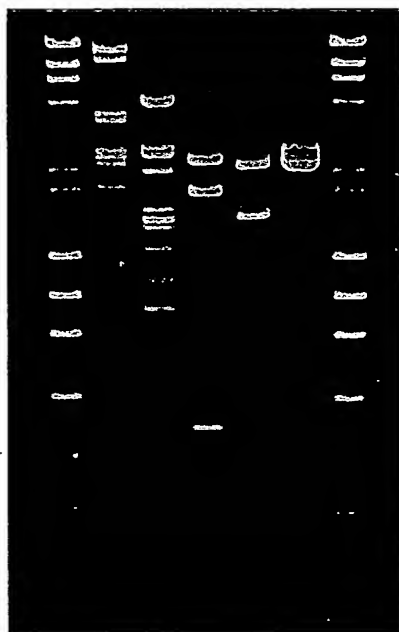# FIG. 1

MmeI DNA CLEAVAGE

FIG. 2A

```
   1 GAATTCCAGA TAGGTAGTCC TTTGGTACTT CCATCCCAAC CAGTGTCACG

  51 TTCCGCGCCA AACCAATCGG TTAAAGTGTA AGAAAGTCTT GCACTGAAGT

 101 AGCTGTAGGA CAAACCGAAG TTAACCTCTG TGGTATCCCA GCGACCACCT

 151 TTAGGTGTTT GACGGAAGCC TGCTGCGTCA CCTGCCAAGT TATATTTCTT

 201 CCATGAACCA CCTGGGTACA GGTAGCTGAT CAAACCAGCA GTCCAACCCA

 251 AGCCTTCAAT AGCAGGAATA GTTCCGTTAT ACCCACCATA AATATCAATT

 301 TCGGCAGTTG CATCAGGGAA GGTATTTGGT GTCACGTTTG AACCCCATGC

 351 ACCGACATAA AAGCCGCTGT CATGAGTAAT ATCAATACCG CCTTGAACGG

 401 CAGGTTTGTG CCAGTTTTGT GAAATACCAC GAGCATAGTA ATCTGAAACA

 451 AATCCAACGT TTGCAGTAGC AGCCCAGGCT GATTTTTCTT CTTTAGCCTC

 501 TTCAGCTGCG TATGAAACTT GGGCAAAAGA TAATGTGCTT AACACTGCTG

 551 TGAGCAATAT AGATTGACGC ATTATGAGTC CTCTCTCTGT GAAATCTTTG

 601 ATTAAGTTGT TGTAAACGAG AATGAAACAA CAACCACAAA GCAAAGCACG

 651 TGCCAAACTA TAAATAACAT TATAATCAAT TATTTAAAAT ATATTTATAA

 701 TCTAAAATAT TAAATTAATT ATTTAATAAA CTGTTTTTTA TTGATTTAAC

 751 TCTAAAACAT ATGGGTGCAA CCACCCTTTT TACTCACTGA TAATGCTAAN

 801 ATAGCCAACA AAGGAGCCTT CACCATGCTG ATTTCAAATG AAAAAATTCA

 851 GGAATTATCT TTAAAAATCA AACAACTAAT CGAATCAAGC CCCATTTCAG

 901 AGCTAAATAA CAACTTGCAT GCACTAATTC AGGGCGCACT CACCAAAATG

 951 GAACTTGTTT CGCGTGAAGA ATTCGATATC CAATCTGCAT TATTAGCGCG

1001 CACGCAAGAG CAATTAAAAC GTCTTGAAGA AAAAATCAGC CAGCTTGAAG

1051 AAGGGCAGGC ATCCAGAAAG TAAAAATTAA TTTACAATTG TTAGCATTCC

1101 ATTATTGAGG AGTGCGCTAT GAGTCTGGCG GTGTTATACA GTCGCGCGTT

1151 AAGCGGCATG GAGGCGCCAG AAGTGGTGGT AGAAGTCCAC TTGGCGAATG

1201 GACTACCCAG CTTTACCATT GTTGAAACAT ATTGAAACTT TAAGCCTTAG

1251 CATTTTTTCA AATATACAAA TGCCCCAAGC TGGTGCATTA AGAAGAATGT
```

FIG. 2B

```
1301  AACAACTCCC TGCAGACTAG GAATAACTTC ATGATTTAAC GAACATCCCT
1351  GAGTTTCAAA GTCGAATCTT CTCGTGTTGC AAATTTCTAC AGCTTCCTTT
1401  CTGACCCTCT TGCACCAAAT TGCACTATGG CGCTAATAAA TCTTCTGCTA
1451  TCCAATAATG TCCAACTAAC CCTTTATGGA CTCTTAAAAA AGATTTAATA
1501  AATGATTAAG ATGAATTCAA GGAATTTGAT GCCTGGAAAT ATGGCAAAAG
1551  CAAAAAGGCA GCCCAGTGCT GACTTTTTTG TTTTAACATT GGCCCATATA
1601  TCCAATTTCA AATAATTTAA AAATTATCGG GAGCTAATCT GTGGCTTTAA
1651  GCTGGAACGA GATAAGAAGA AAAGCTATTG AGTTTTCTAA AAGATGGGAA
1701  GACGCCTCAG ATGAAAACAG TCAAGCCAAA CCCTTTTTAA TAGATTTTTT
1751  CGAAGTTTTT GGAATAACTA ATAAGAGAGT TGCAACATTT GAGCATGCTG
1801  TGAAAAAGTT CGCCAAGGCC CATAAGGAAC AATCTCGAGG ATTCGTAGAT
1851  TTGTTTTGGC CTGGCATTCT TCTTATTGAA ATGAAAAGCA GAGGTAAAGA
1901  CCTCGACAAA GCGTATGACC AGGCACTTGA TTACTTTTCT GGCATTGCAG
1951  AAAGAGACTT ACCCAGATAC GTTTTAGTTT GCGACTTCCA GCGTTTCAGA
2001  TTAACAGACC TAATAACAAA AGAGTCAGTT GAATTTCTTT TAAAGGACTT
2051  ATACCAAAAT GTGAGGTCTT TTGGTTTTAT AGCTGGTTAT CAAACTCAAG
2101  TAATCAAGCC ACAAGACCCT ATTAATATTA AGGCGGCTGA ACGGATGGGT
2151  AAGCTTCATG ACACCCTGAA GTTGGTTGGA TATGAGGGAC ACGCTTTAGA
2201  ACTTTATCTA GTGCGTTTAC TTTTTTGCTT ATTCGCAGAA GACACAACTA
2251  TTTTTGAGAA AAGTTTATTC CAAGAATATA TCGAGACAAA GACGCTAGAG
2301  GACGGCAGTG ACCTTGCACA TCATATCAAT ACACTTTTTT ATGTTCTCAA
2351  TACCCCAGAA CAAAAAAGAT TAAAGAATCT AGACGAACAC CTTGCTGCAT
2401  TTCCATATAT CAATGGAAAA CTTTTCGAGG AGCCACTTCC GCCAGCTCAG
2451  TTTGATAAAG CAATGAGAGA GGCATTGCTT GACTTGTGCT CATTAGATTG
2501  GAGCAGGATT TCACCAGCAA TATTTGGAAG TTTATTCCAA AGCATTATGG
2551  ATGCTAAAAA GAGAAGAAAT CTTGGGGCAC ACTACACCAG CGAAGCAAAT
2601  ATTCTCAAGT TAATCAAGCC ATTGTTTCTT GACGAGCTCT GGGTAGAGTT
2651  CGAGAAAGTT AAAAATAATA AAAATAAATT ACTAGCGTTC CACAAAAAAC
2701  TAAGAGGACT TACATTTTTC GACCCTGCAT GCGGTTGCGG AAATTTTCTT
```

FIG. 2C

```
2751  GTAATCACAT ACCGAGAACT AAGACTTTTA GAAATTGAAG TGTTAAGAGG

2801  ATTGCATAGA GGTGGTCAAC AAGTTTTGGA TATTGAGCAT CTTATTCAGA

2851  TTAACGTAGA CCAGTTTTTT GGTATCGAAA TAGAGGAGTT TCCCGCACAG

2901  ATTGCTCAGG TTGCTCTCTG GCTTACAGAC CACCAAATGA ATATGAAAAT

2951  TTCAGATGAG TTTGGAAACT ACTTTGCCCG TATCCCACTA AAATCTACTC

3001  CTCACATTTT GAATGCTAAT GCTTTACAGA TTGATTGGAA CGATGTTTTA

3051  GAGGCTAAAA AATGTTGCTT CATATTAGGA AATCCTCCAT TTGTTGGTAA

3101  AAGTAAACAA ACACCGGGAC AAAAAGCGGA TTTACTATCT GTTTTTGGAA

3151  ATCTTAAATC CGCTTCAGAC TTAGACCTAG TTGCTGCTTG GTATCCCAAA

3201  GCAGCACATT ACATTCAAAC AAATGCAAAC ATACGCTGTG CATTTGTCTC

3251  AACGAATAGT ATTACTCAAG GTGAGCAAGT ATCGTTGCTT TGGCCGCTTC

3301  TGCTCTCATT AGGCATAAAA ATAAACTTTG CTCACAGAAC TTTCAGCTGG

3351  ACAAATGAGG CGTCAGGAGT AGCGGCGGTT CACTGCGTAA TTATCGGATT

3401  TGGGTTGAAG GATTCAGATG AAAAAATAAT CTATGAGTAT GAAAGTATTA

3451  ATGGAGAACC ATTAGCTATT AAGGCAAAAA ATATTAATCC ATATTTGAGA

3501  GACGGGGTGG ATGTGATTGC CTGCAAGCGT CAGCAGCCAA TCTCAAAATT

3551  ACCAAGCATG CGTTATGGCA ACAAACCAAC AGATGATGGA AATTTCCTAT

3601  TTACTGACGA AGAAAAAAAC CAATTTATTA CAAATGAGCC ATCTTCCGAA

3651  AAATACTTCA GACGGTTTGT GGGCGGGGAT GAGTTCATAA ACAATACAAG

3701  TCGATGGTGT TTATGGCTTG ACGGTGCTGA CATTTCAGAA ATACGAGCGA

3751  TGCCTTTGGT CTTGGCTAGG ATAAAAAAAG TCCAAGAATT CAGATTAAAA

3801  AGCTCGGCCA AACCAACTCG ACAAAGTGCT TCGACACCAA TGAAGTTCTT

3851  TTATATATCT CAGCCGGATA CGGACTATCT GTTGATACCT GAAACATCAT

3901  CTGAAAACAG ACAATTTATT CCAATTGGTT TTGTTGATAG AAATGTCATT

3951  TCAAGTAACG CAACGTATCA TATTCCTAGT GCTGAACCTT TGATATTTGG

4001  CCTGCTTTCA TCGACCATGC ACAACTGCTG GATGAGAAAT GTAGGAGGAA

4051  GGTTAGAAAG TCGTTATAGA TATTCTGCCA GCCTGGTTTA CAACACGTTT

4101  CCATGGATTC AACCCAACGA AAAACAATCG AAAGCGATAG AAGAAGCTGC
```

FIG. 2D

```
4151 ATTTGCGATT TTAAAAGCTA GAAGCAATTA TCCAAACGAA AGTTTAGCTG

4201 GTTTATACGA CCCAAAAACA ATGCCTAGTG AGCTTCTTAA AGCACATCAA

4251 AAACTTGATA AGGCTGTGGA TTCTGTCTAT GGATTTAAAG GACCAAACAC

4301 AGAAATTGCT CGAATAGCTT TTTTGTTTGA AACATACCAA AAGATGACTT

4351 CACTCTTACC ACCAGAAAAA GAAATTAAGA AATCTAAGGG CAAAAATTAA

4401 TTAATGTATT TAACATTAAA CCACCCTGAT TTATTTCGAA TAGTTCAAAT

4451 GCTTCCATGT GGACTAATCG CCTTCAATCA TATTAAAAAA CCGACGCTAG

4501 TAATAAAAAC TTCCAAAGAG GCCATATTAA CCGCCAAAAT TAATCGTGAA

4551 TTTAAAATAT ATCTTTATCA AACCACATCG GCTTGTGTTC TAGTAAGTGC

4601 ATTTTTTGAC GATTCTGATA GTCCACTATT CATTACAACA CCAATTGTTC

4651 GAGATGACCA ACACTCCTTA GACTTGTTAA GATTTTTAAT CAACAATGAT

4701 TTTACGATTT GCTTCTTTGA TGAACTGAAC CGAGAATTTC TTTCCGTTAA

4751 CGCAACTGGT AATTTAGTCT CTATCTTTGA GAGCATTCAC TTGATGCCAC

4801 TGCCGAGCCC AGAGGAAGCC CACAATGCAT TGAATGAAGC GGAATTTTGG

4851 TTCAGTTTAC GCTCAGCTGC TGATGATGAA TCATCTATCC AGGTTTCTTT

4901 ATTGGATAAT CTATTTCCTG ACGATTTTGT AATTTATGAC CTATCCTCAA

4951 ACAAAAACGA TATGACATCA TTGGTTAGAG AAACTAAACC AGGATACTAT

5001 CAGGAAGCAG ATATTGCAAA GTTACTAACA AGAGCTTTTA GTTTGGAAAG

5051 CATTTATCAG AATCCAGTGA AAACAAGCGA TTCAAAAGAG TTGGCAGACG

5101 TTGTGGTATT CGGCCAAAAG GAAATTTTAA TAATTCAAGC TAAAGATAGT

5151 GAAAACAATC AGAAACAAGT TTTAGAGGTT TCGTTAGACA AGAAATGCGC

5201 AAAGTCTTCA AAGAAACTTT CTGAAGCTTT GGCACAACTC ACCGACACTA

5251 TCTTAACAAT ATCCAATACA CCAATAGTTG ATGTTCGGGT TGGTAAGAAA

5301 AAATGCACTC TGAACTTTGA GGGAAAGCAG CTTATTGGTA TCGTCGTTGT

5351 TAAAGAGCTT TTTAATGATA TTTACGATAA ATACAGTCAA AAAGTTTTTG

5401 AGCATGTAGA GTTGTCTAAA GCACCCATTG TCTTCTTTGA CTATCCAGAA

5451 TTTGCAAGAA TGACATTTCA TTGTAATTCT GAGGAATTAT TACTTTATGC

5501 TTTGCATAGG ATATTTAGTT CTGCAATAGA AAATGGAATG TATAAACGAT

5551 TGAGATTTAC TCAACCTATC ATAACTGATG GTCATGACAG CTACTTCAGG
```

FIG. 2E

```
5601  ATACAAAACA GGCCCCATTC TGATGAGGCC TATTTAATTT GCACAGAGGA

5651  TGAAATGAAG CTCTCAAATA AGTTTAAAGA CTAAATTTAT ATTTTCCTCA

5701  GTATCTTAAA AACAATATTC ATTAAATTGG AAAGCCCGCA ATGATTGTTG

5751  CAGTATCAAT GCGGGCATCA GTATCCAGCT CTTGCAATAC ACGGAAGTAT

5801  CAAGAAGCGA ATCAGGATTC TAACCATACC TTTTTAATTG CAACAATCTA

5851  ATTTCCATAA CATGTGTAGC TACATCGAAA AAAAGACCTC GAAGAGGTTG

5901  CAAGAGCGTC CAGCTCGCGG CATCAAAAGA CCCTAGTCTT TTGACAAGGG

5951  GGAGCCAAAA AACTGAGGTG GAGGAGCTTG CCGACGAAGC CAGGAAGCCC

6001  CAGCGTCCGG
```

FIG. 3

```
  1  MALSWNEIRR KAIEFSKRWE DASDENSQAK PFLIDFFEVF GITNKRVATF
 51  EHAVKKFAKA HKEQSRGFVD LFWPGILLIE MKSRGKDLDK AYDQALDYFS
101  GIAERDLPRY VLVCDFQRFR LTDLITKESV EFLLKDLYQN VRSFGFIAGY
151  QTQVIKPQDP INIKAAERMG KLHDTLKLVG YEGHALELYL VRLLFCLFAE
201  DTTIFEKSLF QEYIETKTLE DGSDLAHHIN TLFYVLNTPE QKRLKNLDEH
251  LAAFPYINGK LFEEPLPPAQ FDKAMREALL DLCSLDWSRI SPAIFGSLFQ
301  SIMDAKKRRN LGAHYTSEAN ILKLIKPLFL DELWVEFEKV KNNKNKLLAF
351  HKKLRGLTFF DPACGCGNFL VITYRELRLL EIEVLRGLHR GGQQVLDIEH
401  LIQINVDQFF GIEIEEFPAQ IAQVALWLTD HQMNMKISDE FGNYFARIPL
451  KSTPHILNAN ALQIDWNDVL EAKKCCFILG NPPFVGKSKQ TPGQKADLLS
501  VFGNLKSASD LDLVAAWYPK AAHYIQTNAN IRCAFVSTNS ITQGEQVSLL
551  WPLLLSLGIK INFAHRTFSW TNEASGVAAV HCVIIGFGLK DSDEKIIYEY
601  ESINGEPLAI KAKNINPYLR DGVDVIACKR QQPISKLPSM RYGNKPTDDG
651  NFLFTDEEKN QFITNEPSSE KYFRRFVGGD EFINNTSRWC LWLDGADISE
701  IRAMPLVLAR IKKVQEFRLK SSAKPTRQSA STPMKFFYIS QPDTDYLLIP
751  ETSSENRQFI PIGFVDRNVI SSNATYHIPS AEPLIFGLLS STMHNCWMRN
801  VGGRLESRYR YSASLVYNTF PWIQPNEKQS KAIEEAAFAI LKARSNYPNE
851  SLAGLYDPKT MPSELLKAHQ KLDKAVDSVY GFKGPNTEIA RIAFLFETYQ
901  KMTSLLPPEK EIKKSKGKN*
```

FIG. 4

pTBMmeI.1 IS RESISTANT TO MmeI CLEAVAGE
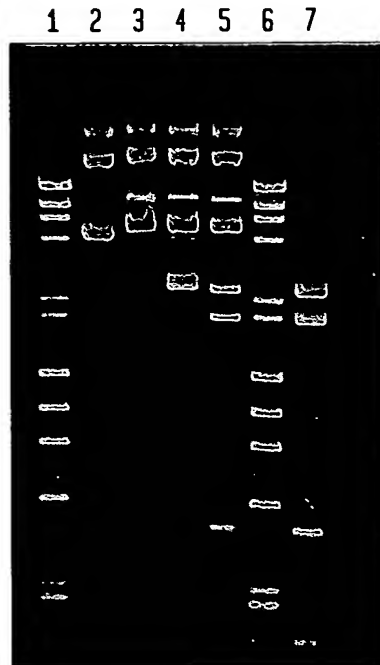
1 2 3 4 5 6 7

FIG. 5

MmeI CLEAVAGE OF HEMI-METHYLATED SUBSTRATES

1 2 3 4 5 6 7 8 9 10 11 12 13 14
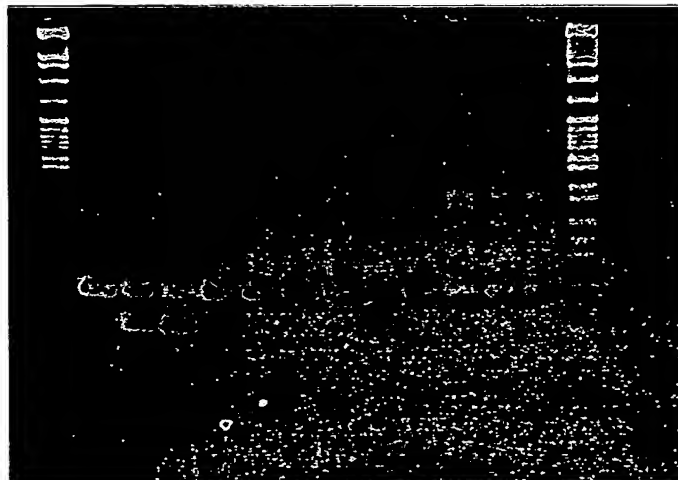
FIG. 6

Top Strand: 5'-TCCG$\underline{A}$C-3'	/	Bottom strand: 5'-GTCGG$\underline{A}$-3'	^{3}H-COUNTS
unmethylated:	/	unmethylated	19,972
unmethylated:	/	methylated:	14,447
methylated:	/	unmethylated:	1,266
methylated:	/	methylated:	917

'A' indicates position of N6-methyl adenine in the DNA substrate

FIG. 7A

PileUp of: @mme.list2

Symbol comparison table: GenRunData:blosum62.cmp   CompCheck: 1102

GapWeight: 6
GapLengthWeight: 1

Name: mmelfeLORF3P = gi|28373198|ref|NP_783835.1| (SEQ ID NO:3)
Name: mmeLre121P = gi|27450519|gb|AA014619.1|AF465251_62 (SEQ ID NO:4)
Name: mme = MmeI amino acid sequence (SEQ ID NO:5)
Name: mmeNMA1791 = gi| 15794682|ref|NP_284504.1| (SEQ ID NO:6)
Name: mmeBSU0677 = gi| 16077744|ref|NP_388558.1| (SEQ ID NO:7)
Name: mmegcry = gi|9945797|gb|AAG03371.1| (SEQ ID NO:8)
Name: mmePf1Q8 = gi|23451826|gb|AAN32074.1|AF461726_1 (SEQ ID NO:9)
Name: saro3834 = gi|23110638|gb|ZP_00096791.1| (SEQ ID NO:10)
Name: mmeMSI135 = gi|20803963|emb|CAD31540.1| (SEQ ID NO:11)
Name: mmeCC0826 = gi| 16125079|ref|NP_419643.1| (SEQ ID NO:12)
Name: mmeDR0119.1 = gi|15807788|ref|NP_285443.1| (SEQ ID NO:13)
Name: mmeDR2267 = gi|15807258|ref|NP_295988.1| (SEQ ID NO:14)

```
              1                                                50
mmelfeLORF3P  ~~~~~~~~~~ ~~~~~~~MPT RQQAAREFVK TWS.SDKKGR EDADRQTFWN
  mmeLre121P  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
         mme  ~~~~~~~~~~ ~~~MALSWNE IRRKAIEFSK RWE.DASD.. ENSQAKPFLI
  mmeNMA1791  ~~~~~~~~~~ ~~~~MKTLLQ LQTAAQNFAA YYK.DQTD.. ERREKDTF*N
  mmeBSU0677  ~~~~~~~~~~ ~~~~~MALID LEDKIAEIVN R.E.DHSD.. .......FLY
     mmegcry  MVMAPTTVFD RATIRHNLTE FKLRWLDRIK QWEAENRPAT ESSHDQQFWG
    mmePf1Q8  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    saro3834  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   mmeMSI135  ~~~~~~~~~~ ~~MSLGAAGL TPITPAAFIK KWRKSELG.. ERQAAQEHFL
   mmeCC0826  ~~~~~~~~~~ ~~~~~~~~~~ ~~MTPAQFVK KWSDSQLR.. ERQASQEHFL
 mmeDR0119.1  ~~~~~~~~~~ ~~~~~~~~~~ ~~MHPQEFAD TWSRRALKAT ERDSYVQHWL
   mmeDR2267  ~~~~~~~~~~ ~~~~~~MPQTE TAQRMEDFVA YW..RTLKGD EKGESQV.FL

              51                                               100
mmelfeLORF3P  DLLQRVYGID N.YYDYITYE KDVQVKADGK VTTRRIDGYI P.STKIMVEM
  mmeLre121P  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
         mme  DFFE.VFGIT N.KR...VAT FEHAVKKFAK AHKEQSRGFV DL...FWPGI
  mmeNMA1791  EFFA.IFGID R.KN...VAH FEYPVKD..P ADNTQ...FV DI...FWEGI
  mmeBSU0677  ELLG.VYDVP R.AT...ITR ....LKK.GN QNLTKRVGEV HLKNKVW...
     mmegcry  DLLDC.FGV. N.ARDLYLY. .....QRSAK RASTGRTGKI DM...FMPGK
    mmePf1Q8  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    saro3834  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~M
   mmeMSI135  D.ICSLVGHP SP.SDEDPTG AFFAFEKGAN KLG.GGKGFA D.VWK..KGH
   mmeCC0826  D.LCRMLEVP TP.AEDDPLG ERYCFERGAA KTG.GGDGWA D.VWR..KGC
 mmeDR0119.1  D.LCQLLHHE APGADPD... ..YKFERRVT KVGTKDKGFA D.VFK..KAH
   mmeDR2267  DRLFQAFGH. ...AGYKEAG AE..LEYRVA KQG.GGKKFA DLLWR..PRV

              101                                              150
mmelfeLORF3P  KGKNIKDLSK PITQSGGD.. .......ELT PFEQAKRYAN FLPN...SEQ
  mmeLre121P  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
         mme  LLIEMKSRGK DL:....D.. .......KAY D..QALDYFS GIAERD...L
  mmeNMA1791  FLAEHKSANK NL.....T.. .......KAK E..QAERYLQ EIGRTKPSAL
```

FIG. 7B

```
mmeBSU0677   .FKEAK.KGK  LF.....D..  .......ALI  DIEQQVEYL.  ........SAK
   mmegcry   VIGEAKSLGV  PLDDA.....  ..........  .YAQALDYLL  G.GTIANSHM
   mmePf1Q8   ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~
   saro3834   NPVEIEEAVS  DLARAPYDAS  EFPFQFLAAF  GNKQTTLQRL  RAGNSNQSDL
 mmeMSI135   FAWEYKRKKG  NLDEA.....  ..........  .LLQLMRYAP  AL........
 mmeCC0826   FGWEYKGKHK  NLDAA.....  ..........  .LRQLQAYAL  DL........
mmeDR0119.1  FITEYKRPGS  DLGAA.....  ..........  .LQQATLYSR  DL........
 mmeDR2267   LI.EMKKRGE  KLANH.....  ..........  .YQQAFDYWL  KL........

              151                                                   200
mmelfeLORF3P  PR........  WILVSNFNEI  DIHDM..E..  .RPLDEPKVI  KL........
 mmeLre121P   ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~
        mme   PR........  YVLVCDFQRF  RLTDLITK..  .ESVE....F  LL........
 mmeNMA1791   PE........  YYAVSDFAHF  HLYRRVPE..  .EGAENQWQF  PL........
 mmeBSU0677   PR........  YLLVTDYDGV  LAKDTKTL..  .EALDVKF..  ..........
   mmegcry    PA........  YVVCSNFETL  RVTRLNRTYV  GDSADWDITF  PL........
   mmePf1Q8   ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~
   saro3834   PGAVLQRNHI  HIATCDAGNV  DRTLAALRKS  PKTASQKARF  ILATDGVAFQ
 mmeMSI135    ....L..SPP  LHIVCDIERL  RIHTAWTNTV  PSTY..VITL  DDLAE.....
 mmeCC0826    ....Q..NPP  YLVVSDMERI  IVHTNWTNTI  SRKI..EFTL  DDLHE.....
mmeDR0119.1   ....G..NPP  LLLTSDFQRI  EINTAFTGTS  PKSY..LITL  DDIAENRVVG
 mmeDR2267    ....VPDRPR  YAVLCNFDEL  .....WVYDF  NQQ......L  DEPMDRLRI.

              201                                                   250
mmelfeLORF3P  .EDLPKKVKS  L........E  F.....MVDA  NQQQVIDEKQ  LSVDAGNLVA
 mmeLre121P   ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~
        mme   .KDLYQNV.R  S.........  FGFIA...GY  QTQVIKPQDP  INIKAAERMG
 mmeNMA1791   .EELPEYITR  G........V  FDFMF...GI  EAKVRQIQEE  ANIQAAATIG
 mmeBSU0677   .EELPQY...  ..........  FDFFLAWKGI  EKVEFEKENP  ADIKAAERFA
   mmegcry    .AEIDEHIEQ  L........A  F.....LADY  ETSAYREEEK  ASLEASRLMV
   mmePf1Q8   ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~
   saro3834   AEDMASGETV  ACNYAAFPDK  FAFFLPLAGI  TTVQQIRESS  FDIKATGRLN
 mmeMSI135    ....PSAREM  LHNVFFSPEK  .........L  RPTR..TRAA  VTKEAADKFS
 mmeCC0826    ....PEKLAM  LRQVFDGSDS  .........L  KPKI..SPQE  LTAKVAQRFG
mmeDR0119.1   GNDVP.ALQI  LHSALHQPYD  .........L  DPRL..FRER  ITTDATRQVG
 mmeDR2267    .EELPERYTV  LNFMFEQ..E  .........R  APLFGNNRVD  VTREAADSVA

              251                                                   300
mmelfeLORF3P  KIYNELTNAY  AAGRGIDVN.  ...EPRIQRS  LN..MLIVRL  VFLLYADDSN
 mmeLre121P   ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~
        mme   KLHDTL.K..  ......LVG.  ....YE.GHA  LE..LYLVRL  LFCLFAEDTT
 mmeNMA1791   RLHDAL.K..  ......EEG.  ...IYE.EHE  LR..LFITRL  LFLFFADDSA
 mmeBSU0677   RIYDVLRK..  ......ENN.  ...IIETNRG  LD..LFLIRL  LFCFFAEDTD
   mmegcry    ELFRAMNGDD  VDEAVGDDAP  TTPEEEDERV  MRTSIYLTRI  LFLLFGDDAG
   mmePf1Q8   ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~
   saro3834   KLYVELLKDN  PDWA......  .......SRS  EDMNHFMARL  IFCFFAEDTD
 mmeMSI135    AIALRVQGR.  G.TPD.....  ..........  .EIAHFVNQL  VFCFFAQSVS
 mmeCC0826    DLGRRLQER.  GHHPR.....  ..........  .DVAHFLNRV  VFCMFAEDAK
mmeDR0119.1   LVARRLGERE  GRT.......  ..........  .RAAHMMMRV  VFALFAEDTG
 mmeDR2267    KVLNSVIAR.  GEDRA.....  ..........  .RAQRFLLQC  VMAMFAEDFE

              301                                                   350
mmelfeLORF3P  LFGKEDIFQA  FIER...REP  RDIRRDLSEL  FKVLDQP.EE  QRDPYLDDEF
 mmeLre121P   ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~
        mme   IFEKS.LFQE  YIETKTLEDG  SDLAHHINTL  FYVLNTP.EQ  KRLKNLDEHL
 mmeNMA1791   VFRRNYLFQD  FLE..NCKEA  DTLGDKLNQL  FEFLNTP.DQ  KRSKTQSEKF
 mmeBSU0677   IFKRNS.FTN  LIKTLTEEDG  SNLNKLFADL  FIVL....DK  NERDDVPSYL
```

FIG. 7C

```
      mmegcry  LWDTPHLFAD FVRNETTPE. .SLGPQLNEL FSVLNTA.PE KRPKRLPSTL
     mmePf1Q8  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     saro3834  IFVGEGLFSR TVETMSARDA SDTHMVIAEI FRAMDTRLAD RAAAGIKSWA
    mmeMSI135  LLPD.GLFTK LLK.RSARAP ERAMSYLDKL FEAME..... RGGEF...DL
    mmeCC0826  LLPE.GLFTR LTRSMQMRPP AEAAPQFDAL FAMMR..... AGGMF...GA
   mmeDR0119.1 MLER.GIVTR LLE.RARAPP GEDQLYFQDL FGAMK..... GGGEF...WG
     mmeDR2267 LIPR.GFFTE LADD.ARAGR GSSFDLFGGL FRQMNTSERA RGGRF.....

                351                                             400
  mmelfeLORF3P  NQFAYVNGGM FSDENVIIPQ FTDELKRLIV EDAGRGFDWS GISPTIFGAV
   mmeLre121P   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
          mme   AAFPYINGKL FEEPLPPA.Q FDKAMREALL .DLCS.LDWS RISPAIFGSL
   mmeNMA1791   KGFEYVNGGL FKERLRTF.D FTAKQHRALI .DCGN.FDWR NISPEIFGTL
   mmeBSU0677   KEFPYVNGQL FTEPHTEL.E FSAKSRKLII .ECGELLNWA KINPDIFGSM
      mmegcry   AKFPYVNGAL FAEPLAS.EY FDYQMREALL AAC..DFDWS TIDVSVFGSL
     mmePf1Q8   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     saro3834   DVFPYVNGQL FSGS.TECPR FSKIARSYLL H..IGSLDWQ KINPDIFGSM
    mmeMSI135   TDITWFNGGL FDGR..RALR LDDGDIGLL. .VAADSLDWG LIDPTIFGTL
    mmeCC0826   DIVHWFNGGL FDEK..PALP LERADIKLIH DTAAEH.DWS DLDPSVFGNM
   mmeDR0119.1  TDIRHFNGGL FDSE..DALA LTSEDAAAL. .IIAAKLDWS EVEPSIFGTL
    mmeDR2267   APIPYFNGGL FRAV..DPIE LNRDELYLLH KAALEN.NWA RIQPQIFGVL

                401                                             450
  mmelfeLORF3P  FESTLN.PET RRSGGMHYTS IENIHKVIDP LFLNDLHDEF D.........
   mmeLre121P   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
          mme   FQSIMD.AKK RRNLGAHYTS EANILKLIKP LFLDELWVEF E.........
   mmeNMA1791   FQSVMD.AQE RREAGAHYTE AANIDKVING LFLENLRAEF E.........
   mmeBSU0677   IQAVAS.EES RSYLGMHYTS VPNIMKVIKP LFLDKLNQSF ..........
      mmegcry   FQLVKS.KEA RRSDGEHYTS KANIMKTIGP LFLDELRAEA D.........
     mmePf1Q8   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     saro3834   IQAVAD.DEE RGALGMHYTS VPNILKVLNP LFLDDLRAKL E.........
    mmeMSI135   FERFLD.PEK RAQIGAHYTD PEKIMRLVDP VILRPLRQEW EQARREIVEL
    mmeCC0826   FEEALKATRE RAALGAHYTD REKILKIIDP VITWPLMAQW ETALAEIRAA
   mmeDR0119.1  FENSLDV.DT RSRRGAHYTS VNDIERIVDR VVMEPLWAEW D.........
    mmeDR2267   FQSSMDKKEQ HAK.GAHYTS EADIMRVVLP TIVTPFQRQI EAATTQ....

                451                                             500
  mmelfeLORF3P  .......... .......... .......... .......... ....KIQNMG
   mmeLre121P   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
          mme   .......... .......... .......... .......... ....KVK...
   mmeNMA1791   .......... .......... .......... .......... ....AVK..A
   mmeBSU0677   .......... .......... .......... .......... .........L
      mmegcry   .......... .......... .......... .......... ....KL..VS
     mmePf1Q8   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     saro3834   .......... .......... .......... .......... ....E.....
    mmeMSI135   LNGN...... .......... .RKPPMRR. ..QQSRR... ...MKREEAA
    mmeCC0826   LDARAAAEAE RKAVLEAAAE AMRADPVKAK AGEAARRKTL TAIAKRSDAA
   mmeDR0119.1  .......... .......... .......... ....ALRLSL PELKK.....
    mmeDR2267   .......... .......... .......... .......... ..........

                501                                             550
  mmelfeLORF3P  NRRQRVTRAK AFRDKLGKLK FFDPACGSGN FLTETYLSLR KMENECLRII
   mmeLre121P   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
          mme   ...NNKNKLL AFHKKLRGLT FFDPACGCGN FLVITYRELR LLEIEVL.RG
   mmeNMA1791   LKRDKAKKLA AFYQKIQNLQ FLDPACGCGN FLIVAYDRIR ALEDDIIAEA
   mmeBSU0677   DAYDDYTKLE NLLTRIGKIK FFDPACGSGN FLIITYKELR RMEINIIKRL
      mmegcry   SPSTSVAALE RFRDSLSELV FADMACGSGN FLLLAYRELR RIETDIIVAI
     mmePf1Q8   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
```

FIG. 7D

```
      saro3834  .AGDNSRKLL NLRNRMAKIR VFDPACGSGN FLVIAYKQMR ELEAEI....
    mmeMSI135  .AEVR.SR.. .FTERLRKLR ILDPACGSGN FLYLALQGVK DIEHRANLDC
    mmeCC0826  LGQAK.DRLE AFLSRLAAFR VLDPACGSGN FLYVALHALK DIERRALVDA
  mmeDR0119.1  ..NVRLERLF AFQDRLTAVR ILDPACGSGN FLFVALKKLL DLEYQVRMAA
    mmeDR2267  ......KELR AILDELASFQ VLDPACGSGN FLYVAYRELR RLEARALL..
```

```
                551                                              600
mmelfeLORF3P  VGNQGA..LA LTDESEPKVK IQNFYGIEIN DFAVSVARTA MWIAESQMWE
  mmeLre121P  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
         mme  LHRGGQ..QV LDIEHLIQIN VDQFFGIEIE EFPAQIAQVA LWLTDHQMNM
  mmeNMA1791  LKDKAD..GL FD.SPSVQCR LKQFHGIEID EFAVLIARTA MWLKNHQCNI
  mmeBSU0677  QELLGE..YL Y..VPSV..T LSQFYGIEIE DFAHDVAKLS LWIAEHQMNE
     mmegcry  RQRRGETGMS LNIEWEQKLS IGQFYGIELN WWPAKIAETA MFLVDHQANK
     mmePf1Q8  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    saro3834  NRRRGEADRR SD......IP LTNFRGIELR NFPAEIARLA LIIAEYQCDV
   mmeMSI135  ..EM..LG.M PAQLP..LVG PEILRGIEIN MMAAELARTT IWIG.DIQWQ
   mmeCC0826  ..ER..LG.L EVPTP..RVG LACVRGIEIE EYAAELARVT LWIG.DLQWH
 mmeDR0119.1  ..VMNDIGEF EMP.P..LVH PQQMLGIEIE TFAHELASIT LWMG.YFQWK
   mmeDR2267  ..RLRDLSAP GTALPPARVS IRQMHGLEYD PFGVELAKVT LTLAKELAIR
```

```
                601                                              650
mmelfeLORF3P  QTKDI..TFA NKDFLP.LDS NDSIYEGNAL RMDWNDI... ......VKPY
  mmeLre121P  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
         mme  KISDEFGNYF A..RIP.LKS TPHILNANAL QIDWNDV... ......LEAK
  mmeNMA1791  RTQIRFDGEV ACHTLP.LED AAEIIHANSL RTPW...... ......QAA
  mmeBSU0677  ELKNEVHNAV R.PTLP.LHT AGDIRCANAI RVEWTEVCP. ......AQGS
     mmegcry  ELANAVGR.. PPERLP.IKI TAHIVHGNAL QLDWADILS. ......ASAA
     mmePf1Q8  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    saro3834  LYRGQ..KEA LAEFLP.LDS QNWITCGNAL RLDWLSICPP TGTAVKLQAN
   mmeMSI135  IKNGIRS... ..KSIPILRK LDAIERRDAL VRQAQDVDTA RDAQG.....
   mmeCC0826  AKNNYRG... ..FAEPILSS LDQIECRDAL L......... .NADG.....
 mmeDR0119.1  RAHG.GH... ..WETPILQR LDNIQNRDAL L......... .NPDG.....
   mmeDR2267  EMHDLLGNTG LDFDQPL..P LDNLD..DRI V......... ...QG.....
```

```
                651                                              700
mmelfeLORF3P  EL........ .......... DYIMGNPPFV GYSLQTKEQK QDIKQEFFKY
  mmeLre121P  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
         mme  KC........ .......... CFILGNPPFV GKSKQTPGQK ADL.LSVFGN
  mmeNMA1791  D......... .......... .YIFGNPPFI GSTYQTKEQK NDL.ESICGH
  mmeBSU0677  EE........ .......... VYVFGNPPYL GSKKQNKEHK SDM.LSIFGK
     mmegcry  K......... .......... TYIFGNPPFL GHATRTAEQA QELR.DLWG.
     mmePf1Q8  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    saro3834  DLFEMPLDQA EIDFENEGGE TYICGNPPYL GAKKKSSDQI EDMKRV...G
   mmeMSI135  DLLAALQPVS EDAEAEWPEA EFIVGNPPFV GVRLMRQALG DPTVDRLFDV
   mmeCC0826  .......... ..TEAQWPAV DVIVGNPPFL GSKRLRDGLG NDYVERLFST
 mmeDR0119.1  .......... ..TEATWPRA DFIVGNPPFL GDKMMRSQLG EAYTTQLRET
   mmeDR2267  DALF...... ....TPWPRV DAIVGNPPFQ SKNKLQREMG AAYVKKLRAH
```

```
                701                                              750
mmelfeLORF3P  TDKY..GKFD YVSGWYIKGA KYIQ.NSTIK VGFVSTDSII QGEQAPEIWK
  mmeLre121P  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
         mme  LKSA..SDLD LVAAWYPKAA HYIQTNANIR CAFVSTNSIT QGEQVSLLWP
  mmeNMA1791  IKGY..GLLD YVCNWYVKAA GIMAQHPQVQ TAFVSTNSIC QGQQVEILWG
  mmeBSU0677  VKNG..KMLD YISAWFYFGA KYAST.TNAK VAFVSTNSVT QGEQVSILWN
     mmegcry  TKDI..SRLD YVTGWHAKCL DFFKSREG.R FAFVTTNSIT QGDQVPRLFG
     mmePf1Q8  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    saro3834  LDKA..QLLD YVSAFIVRGL PLVAQQ.RCD MALVSTSSIC QGEQVSLIWP
```

FIG. 7E

```
mmeMSI135       YDGRVSREAD LVCYWVEKSR AAVAADRTRR VGLVTTNSIR GGANRR.VLD
mmeCC0826       YRGKVPAEAD FVAYWIAKAW ELVQAQQGRR AGLVTTNSVR GGASRK.VLD
mmeDRO119.1     FKDRLPGQSD LVCYWPEKAR ALIEAGVTTR AGFVTTNSIR GGKNRV.VLE
mmeDR2267       YPD.VPGRAD YCVYWIRKAH D..QLGSGQR AGLVGTNTIR QNDSRVGGLD

                751                                                    800
mmelfeLORF3P    VLFNDFHIFI NYGYRSFEWN NEAANKAKVD VVIVGFSTK. .EDKNPTIYD
mmeLre121P      ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~VNKDKILYN
mme             LLLS.LGIKI NFAHRTFSWT NEASGVAAVH CVIIGFGLKD ..SDEKIIYE
mmeNMA1791      SLLN.QGIEI HFAHRTFQWT SQAAGKAAVH CIIVGFRQKP PMPSEKTLYD
mmeBSU0677      ELFK.FGIQI NFAYKSFKWA NNAKNNAAVI VVIVGFG... PLDTKVNKYL
mmegcry         PIFKA.GWRI RFAHRTFAWD SEAPGKAAVH CVIVGFDKES ..QPRPRLWD
mmePf1Q8        ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
saro3834        RILKSAN..V KFAYRPFRWS NSAANNAGVY CTIIGLTGSE VSNKK.....
mmeMSI135       RIIA.ES.RL FEAWSDEPWV VDG...AAVR VSLICFGHG. .EDPLCL...
mmeCC0826       PIAD.AG.AL MEAWADEPWA LEG...AAVR VSMFGFGDG. .FAERRL...
mmeDRO119.1     RIKA.TG.DL FMAWPDEPWQ QNG...AAVR VSLFGFDNG. .TETLRT...
mmeDR2267       YVVQHGG.TI TDAVGTQVWS GD.....AAVH VSIVNWVKGP AEGPKHLAWQ

                801                                                    850
mmelfeLORF3P    EQKIIS.... ..A.KHINQY MYDSDNIFID TTRKY.IEA. MPKMKTGNRP
mmeLre121P      SSN*ISH... ..C.KNINGY LFDGNNIFV. TNRPAPLSN. VPRMHNGCKL
mme             YESINGEPLA IKA.KNINPY LRDGVDVIA. CKRQQPISK. LPSMRYGNKP
mmeNMA1791      YPDIKGEPEK HAV.ANINPY LIDAPDLII. AKRSRPIHC. EPDMVNGSKP
mmeBSU0677      FVD...ETKK L.V.SNISPY LTDGENILV. SSRTKPISD. LPKLHFGNMP
mmegcry         YPDVKGEPVS VEVGQSINAY LVDGPNVLVD KSR.HPISSE ISPATFGNMA
mmePf1Q8        ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~P.ADE RSQMDAGGKP
saro3834        ...LFGEGSV VEC.SSIAPY LVPGPEI.IC APRQSSIS.G FARMVMGSNP
mmeMSI135       ..DGRT.... ...VAQINAD LTAGVTDLTK ARRL.S..EN QNVAFMGDTK
mmeCC0826       ..EGRK.... ...AEHLHSD FRGASTDVTK ALRL.K..EN ASIAFMGDTK
mmeDRO119.1     LNDGH..... ...VGVINAD LNAG.TDVKQ AQKL.P..EN AGVSFIGTQK
mmeDR2267       VGDHRTSPWQ STELPVINSA LSAG.TDVTQ AQKL.RVNMN SGACYQGQTH

                851                                                    900
mmelfeLORF3P    ADGGALILSP KEAKEL.VNE EPQSKQ...F IKKLTGSKEF IT.GKYRYCL
mmeLre121P      LDGGFYTLTS QERKEA.ISK DPYADK...F IRPYLGAKNF IH.GTARYCI
mme             TDDGNFLFTD EEKNQF.ITN EPSSEK...Y FRRFVGGDEF IN.NTSRWCL
mmeNMA1791      TEGGNLILST AEKDAL.IAA EPLAEQ...Y IRPFIGADEF LN.GKTRWCL
mmeBSU0677      NDGGGLLFTI TEYTDA.INK YPELVP...Y FKKFIGSVEF IN.GGLRYCL
mmegcry         RDGGNLLVEV DEYDEV.MS. DPVAAK...Y VRPFRGSREL MN.GLDRWCL
mmePf1Q8        VEGGNLLFAE EEKQRL.VEG NVDVVK...F LKRVYGASEY IR.GEVRFCL
saro3834        VDGKRLIFEQ DEKESV.VAA DPRSER...F FKRYGGTQEL VN.GVDRWCL
mmeMSI135       G.GAFDVPGS LARAWLSMPM NPNGRPNSDV LRPWRNGMDV ARRG...RDM
mmeCC0826       G.GAFDVSGE IAREWLRLPL NPNGRPNSDV LKPWRNAMDM TRRS...SDK
mmeDRO119.1     G.GAFDIPGD LARSWLSVP. NPDRVSNADV LKPWVNGMDL TRRP...SGR
mmeDR2267       GHKGFLLDGL EAGQMLSAE. ....RKNAEV IFPYLTGDEL LRTSPPHPTR

                901                                                    950
mmelfeLORF3P    WLVNVTP... .KQLR..SMP L..VLKRVE. QCKENR.LSG APDRQKLAAT
mmeLre121P      WLKDANP... .KDIH..QSP F..ILDRIN. KVAEFRSQQK SKDTQKYAKR
mme             WLDG.ADIS. ..EIR..AMP L..VLARIK. KVQEFRLKSS AKPTRQSAST
mmeNMA1791      WFHGVSDVKR NHDLK..QMP Q..VQARIQ. AVKTMREASS DKQTQKDAAT
mmeBSU0677      WLNE....AK YEKIK..SNP L..IQERIS. ISKNHREKST DKGTNKLALT
mmegcry         WLVDVAP... .SDIA..QSP V..LKKRLE. AVKSFRADSK AASTRKMAET
mmePf1Q8        WISD....SQ EQEAK..SNS D..INCKLN. AVAAFRLKSP KAATKKGAAW
saro3834        WIND....DQ VDDAK..AIA E..IAKVLE. SCRSYR.QGA GRDAQKAANR
mmeMSI135       WIVDFGWEMS EQEAALYEAP FQHIREHVFP ........ER SKN.......
```

FIG. 7F

```
mmeCC0826  WIIDFGWTMS EADAALFETP FRHVLLHVKP ........ER DRN.......
mmeDR0119.1 WIIDFA.QMD EGEARQYLQP MAYVEQKIRP ........ER ATN.......
mmeDR2267  YVIDF.QPRD VFGARAYKLP FARIEREVLP TRQAAAAEEE ARNAEVLAAN
```

```
            951                                          1000
mmelfeLORF3P P.......... ........HL ......FREQ MNPDNYMIVP LVTGCRRKYV
mmeLre121P  P.......... ........ML PTRLAYYSHD EHTD.MLIVP ATSSQRREYL
        mme P.......... ........M. ..KFFYISQP .DTDY.LLIP ETSSENRQFI
mmeNMA1791  P.......... ........W. ..LFQKIRQP SDGNY.LIIP SVSSESRRFI
mmeBSU0677  P.......... ........W. ..KFRDTHE. .TTNYSIVVP SVSSENRFYI
     mmegcry P.......... ........HL FGQ....RSQ PDTDY.LCLP KVVSERRSYF
mmePf1Q8    P.......... ........H. ..KFEEVKQI GN.EVVTIVP KVSSESREYL
   saro3834 P.......... ........HS FC.YRTFQE. ...NIGIHVG LTIGNGLSHV
mmeMSI135   ....RRDAYR ...ERWWRHV EPRPAFHASL QGHSRYMATP RV.AKHRTFV
mmeCC0826   ....NREMYR ...LNWWKHV EPRQGLMKRV PALSRLLVTP EV.SKHRLFI
mmeDR0119.1 ....SDRPSR ...ERWWLHQ RSRPELREAT IELDRFIGIP RV.AKHLLPV
mmeDR2267   PKAKTNKHHR NFLNQWWALS YGRSEMIEKI SSLSRYIVCS RV.TKRQVFE
```

```
            1001                                         1050
mmelfeLORF3P PFGYLG.NDI IPTNLATIIP EADHYAFGVL ESIVHMAWMR VVA...GRKG
mmeLre121P  PIGYVSEKNI VSYSL.MLIP NASNFNFGIL ESKVHYIWLK NFC...GRLK
        mme PIGFVDRNVI SS.NATYHIP SAEPLIFGLL SSTMHNCWMR NVG...GRLE
mmeNMA1791  PIGYLSFETV VS.NLAFILP NATLYHFGIL SSTMHNAFMR TVA...GRLK
mmeBSU0677  PMGLAGADTI LS.NLIYVIY DAEIYLLGIL MSRMHMTWVK AVA...GRLK
     mmegcry TVQRYPSNVI AS.DLVFHAQ DPDGLMFALA SSSMFITWQK SIG...GRLK
mmePf1Q8    PVGLLPRGSI VT.DLAFALY DAPLWNMALI ASRLHLVWIG ~~~~~~~~~
   saro3834 PAD.LKSSGF VSSHTAYMIY GWHPVEFALL NSRLMLVWTE TVG...GRLG
mmeMSI135   WLD....QAI VPDSRIFAFS RSDDVFFGIL HSRFHEAWSF GTCSWHGV.G
mmeCC0826   WLD....ARV LPDHKLQVVT LDDDCSFGVL HSRFHEVWAL AAGSWHGS.G
mmeDR0119.1 WLP....EGT LPDSQVVVIA RDDDFIFGVL ASTIHRSWAR MQGTYMGV.G
mmeDR2267   FLD....NGI RPSDGLQIFA FEDDYSFGVI QSSVHWQWLI ARG...GTLT
```

```
            1051                                         1100
mmelfeLORF3P TSYRYSKNLV YTNFPWPV.V DINQKEK... .......ITI TAQDILNARN
mmeLre121P  SDYRYSNTII YNNFPWPT.V GDKQEQN... .......ISE TAQGILNTRK
        mme SRYRYSASLV YNTFPWIQ.. ....PNEKQ. ....SKAIEE AAFAILKARS
mmeNMA1791  SDYRYSNTVV YNNFPFPE.. SCRLPSENDR PDPLRAAVEA AAQTVLDARG
mmeBSU0677  TDYRYSAGLC YNTFPIPE.L STRRKNE... .......IEE AILEILDLR.
     mmegcry SDLRFANTLT WNTFPVPE.L DEKTRQR... .......IIK AGKKVLDARA
mmePf1Q8    ~~~~~~~~~ ~~~~~~~~~ ~~~~~~~~~ ~~~~~~~~~ ~~~~~~~~~
   saro3834 NGMRFSNTIV YNTFPVPS.L TDQNK..... ......ADLTR CAEDILLARE
mmeMSI135   NDPTYNSAGV FETFPFPEGL TPDIPAVRYE KDSRAIAISK AAKRLDDIRN
mmeCC0826   NDPRYTISTT FETFPFPEGL TPNIAAVDYE GDPRAQAIAA AAAELNRLRE
mmeDR0119.1 NDLRYTPSTC FETFPVP... APT........ .DEQRAEIEK WAKYIVQLRE
mmeDR2267   ARLMYTSDTV FDTFPWP*D. .PTLAQVR.. ......AVAA AAVKLRELRN
```

```
            1101                                         1150
mmelfeLORF3P .......... .......... ........LY..... ....PDSSLA DLYDPLTMPI
mmeLre121P  .......... .......... ........LY..... ....PDSSLA DLYDPLTMPV
        mme .......... .......... ........NY..... ....PNESLA GLYDPKTMPS
mmeNMA1791  .......... .......... ...QYRREAQ EAGLPEPTLA ELYAPDAGYT
mmeBSU0677  .......... .......... .......EE QGG....TLA ELYNPSTMPI
     mmegcry .......... .......... ........LH..... ....PERSLA EHYNPLAMAP
mmePf1Q8    ~~~~~~~~~ ~~~~~~~~~ ~~~~~~~~~ ~~~~~~~~~ ~~~~~~~~~
   saro3834 .......... .......... .......SHFP..... ......ATIA DLYDPETMPE
mmeMSI135   AWLNPSDLVQ IKPEVVPGYP DRILPKDIAS DAILRDRTLT NLYNR.....
mmeCC0826   AWLNPPDLVR IEPEVVPGYP DRVLPVSPEA GAELKKRTLT NLYNQ.....
```

FIG. 7G

```
mmeDRO119.1  HLLN...... ..........  .........Q DA...KGTLT GIYNQLEKLR
mmeDR2267    KVMREQ.... ......GW. ..........SLR DLYRTLDMPG

             1151                                           1200
mmelfeLORF3P E......... .LRKAHEAND KAVLKAYGLK P.SATEPEI. .VQHLFKMYE
mmeLre121P   E......... .LRKAHEAND KAVLKAYGLS P.KATEQEI. .VEHLFKMYE
mme          E......... .LLKAHQKLD KAVDSVYGFK .GPNT..EIA RIAFLFETYQ
mmeNMA1791   A......... .LDKAHATLD KAVDKAYGYK TGKNTDDEAE RVAFLFELYR
mmeBSU0677   E......... .LKVAHEKLD GIVERAYRQK QFES..DE.E RLEVLLKLYQ
mmegcry      E......... .LIKAHDALD REVDKAFGAP RKLTTVRQ.. RQELLFANYE
mmePf1Q8     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
saro3834     S......... .LRAAHDRND EVLERIY... IGRRFRNDTE RLEKLFELYT
mmeMSI135    ......RP.Q WLVDAHSDLD AAVAGAYGWP ADIS...EDE ALANLLELNL
mmeCC0826    ......RP.A WLDMAHQRLD AAVAAAYGWP DGLT...DDE ILERLFALNQ
mmeDRO119.1  NSPDAAHPVS ALATAHDKLD QAVATAYGWE WPLN...EDQ VLERLLALNL
mmeDR2267    KNP....... .LRDAQERLD AAVSAAYGLP AGA......D MLDFLLALNA

             1201                                           1250
mmelfeLORF3P KLTKKDW~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
mmeLre121P   KLTKGER*~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
mme          KMTSLLPPEK EIKKSKGKN* ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
mmeNMA1791   KAAAIA~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
mmeBSU0677   EMTER~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
mmegcry      KLISHQP~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
mmePf1Q8     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
saro3834     KMTGGRSSEG GAA~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
mmeMSI135    AREAFNEHAK SGLKTRKPRR RPTPEEVRRA PQMKLPIAGG RKSVVGPQQL
mmeCC0826    ERAAAGR~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
mmeDRO119.1  ERCPA~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
mmeDR2267    XVAAAEARGA AVTGPGLPAG LNTADFVTAD AVRPLG*~~~ ~~~~~~~~~~

             1251                1273
mmelfeLORF3P ~~~~~~~~~~ ~~~~~~~~~~ ~~~
mmeLre121P   ~~~~~~~~~~ ~~~~~~~~~~ ~~~
mme          ~~~~~~~~~~ ~~~~~~~~~~ ~~~
mmeNMA1791   ~~~~~~~~~~ ~~~~~~~~~~ ~~~
mmeBSU0677   ~~~~~~~~~~ ~~~~~~~~~~ ~~~
mmegcry      ~~~~~~~~~~ ~~~~~~~~~~ ~~~
mmePf1Q8     ~~~~~~~~~~ ~~~~~~~~~~ ~~~
saro3834     ~~~~~~~~~~ ~~~~~~~~~~ ~~~
mmeMSI135    TTKDRENQPT SAERPRNTKR RTS
mmeCC0826    ~~~~~~~~~~ ~~~~~~~~~~ ~~~
mmeDRO119.1  ~~~~~~~~~~ ~~~~~~~~~~ ~~~
mmeDR2267    ~~~~~~~~~~ ~~~~~~~~~~ ~~~
```